Immunity, Vol. 7, 473–481, October, 1997, Copyright ©1997 by Cell Press

X-Ray Crystal Structure of HLA-DR4 (DRA*0101, DRB1*0401) Complexed with a Peptide from Human Collagen II

Andréa Dessen,* C. Martin Lawrence,* Susan Cupo,§ Dennis M. Zaller,§ and Don C. Wiley*†‡‖
*Laboratory of Molecular Medicine
†Howard Hughes Medical Institute
The Children's Hospital
Boston, Massachusetts 02115
‡Department of Molecular and Cellular Biology
Harvard University
Cambridge, Massachusetts 02138
§Molecular Immunology
Merck Research Laboratories
Rahway, New Jersey 07065

Summary

Genetic predisposition to rheumatoid arthritis (RA) is linked to the MHC class II allele HLA-DR4. The charge of the amino acid at DRβ71 in the peptide-binding site appears to be critical in discriminating DR molecules linked to increased disease susceptibility. We have determined the 2.5 Å x-ray structure of the DR4 molecule with the strongest linkage to RA (DRB1*0401) complexed with a human collagen II peptide. Details of a predicted salt bridge between lysine DRβ71 and aspartic acid at the P4 peptide position suggest how it may participate in both antigen binding and TCR activation. A model is proposed for the DR4 recognition of collagen II (261–273), an antigen immunodominant in human-transgenic mouse models of RA.

Introduction

Increased susceptibility in humans to specific autoimmune diseases has a strong genetic association with specific class I and class II alleles of the major histocompatibility complex (MHC) (reviewed by Tiwari and Terasaki, 1985; Todd et al., 1988; Wucherpfennig and Strominger, 1995). Genetic predisposition to rheumatoid arthritis (RA) has been strongly associated with the MHC class II HLA-DRB1 locus (Nepom et al., 1989; Lanchbury et al., 1991; Nepom and Erlich, 1991). Approximately 80% of Caucasian patients with RA express DR4 (DRB1*0401, DRB1*0404) or DR1 (DRB1*0101) allotypes (Feldman et al., 1996). DRB1*0401 confers higher risk than any other allele; it is carried by 50% of all patients testing positive for rheumatoid factor (Gregersen et al., 1987; Nepom and Erlich, 1991). Of patients with a particularly severe form of arthritis (Felty's syndrome), 92%–95% express the HLA-DR4 molecule, with increased disease frequency mainly due to the presence of DRB1*0401 (Lanchbury et al., 1991). The presence of DR4 is also linked to increased duration of chronic arthritis in patients with Lyme disease (Steere et al., 1990).

The DRβ chains encoded by the RA-related DRB1 genes possess a "shared epitope" formed by very similar amino acids at positions 67–74 (Table 1) (Gregersen et al., 1987; Nepom et al., 1989; Hiraiwa et al., 1990), now known to form part of the α helix at one edge of the peptide-binding site of DR molecules (Brown et al., 1993). The majority of RA patients in the groups that do not carry the HLA-DR4 subtype (e.g., specific ethnic groups), although they carry alleles with amino acid variation in the peptide-binding site, share identical residues with the DRB1*0401 sequence at the shared epitope (Nepom et al., 1989; Nepom and Erlich, 1991). Site-directed mutagenesis experiments of the 67–74 DR4β region (Hiraiwa et al., 1990; Fu et al., 1995; Signorelli et al., 1995) as well as studies of spontaneous mutants in mice (Mengle-Gaw et al., 1984) indicate that sequence differences in this region, especially in residue β71, can profoundly influence T cell recognition and immune response.

The shared epitope residues are critical in selecting specific amino acids at position 4 (P4) of peptides that will bind to DR4. In particular, among peptides with electrostatically charged residues at P4, only those with negatively charged residues (Asp and Glu) at this position bind to DR molecules with associated increased susceptibility to RA; peptides with such residues do not bind to molecules such as DRB1*0402, an allotype that differs in sequence only at the shared epitope and that is not associated with increased susceptibility to the disease (Table 1) (Hammer et al., 1995; Woulfe et al., 1995; Wucherpfennig and Strominger, 1995). The sequence differences between DRB1*0401 and DRB1*0402 change the electrostatic charge in the pocket that binds the P4 side chain of peptides. Residue DRβ71 is positively charged (Lys or Arg) in DRB1*0401 and other RA-associated allotypes and is negatively charged in the nonassociated DRB1*0402 (Table 1). This suggests that the mechanism by which specific DR molecules are associated with RA could involve their selectivity for binding a peptide with a negatively charged P4 (Hammer et al., 1995; Wucherpfennig and Strominger, 1995). Such peptide selectivity may alter the T cell repertoire during T cell differentiation in the thymus or selectively activate arthritogenic T cells at the level of antigen presentation, leading to the development of autoimmunity (Gregersen et al., 1987). An alternative hypothesis—that the sequences of the shared epitope are themselves presented as a peptide by another MHC molecule—cannot be ruled out (Zanelli et al., 1995, 1996).

The antigen(s) responsible for the induction of the autoimmunity in RA are unknown. Type II collagen (CII) is a candidate antigen since it is the predominant protein of joint cartilage and since auto-antibodies against CII are found in elevated levels in the serum and joints of patients with RA (Banerjee et al., 1988). If CII is not the antigen that induces RA in humans, it seems that it is involved in at least part of the autoimmune reaction that results in joint inflammation. In a number of rodent models, RA can be induced by immunization with CII (Fugger et al., 1994; Khare et al., 1995; Rosloniec et al., 1997). A set of eight peptides that bind selectively to

‖ To whom correspondence should be addressed (e-mail: wiley@xtal0.harvard.edu).

Table 1. Amino Acids in the Shared Epitope

	67	70	71	74	RA Association
DRB1*0401	Leu	Gln	Lys^+	Ala	++
DRB1*0404	—	—	Arg^+	—	++
DRB1*0101	—	—	Arg^+	—	++
DRB1*0402	Ile	Asp^-	Glu^-	—	None

DR molecules associated with RA, when compared to DR molecules not associated with RA, were identified from four candidate antigens in the pathogenesis of arthritis (Hammer et al., 1995). Two of these peptides were from collagen II: CII(1168–1180) and CII(1350–1362). These observations suggested that a structural study of DR4 complexed with a CII peptide could provide insight into the shared epitope and the mechanism of the MHC association with the development of RA.

We have determined the x-ray crystal structure of the HLA-DR4 allotype DRB1*0401, which confers the greatest susceptibility to RA, complexed to the CII(1168–1180) peptide (Hammer et al., 1995). The crystal also contains Staphylococcus aureus enterotoxin B (SEB), which promoted the growth of ordered crystals. The structure, refined to 2.5 Å resolution, reveals that the peptide conformation is extremely similar to that observed in other MHC class II/peptide complexes (Brown et al., 1993; Jardetzky et al., 1994; Stern et al., 1994a; Stern and Wiley, 1994b; Ghosh et al., 1995; Fremont et al., 1996; Jardetzky et al., 1996). For the first time, no dimer of dimers was observed in an MHC class II crystal. In this article we also discuss the structural basis for the selective binding of peptides to DRB1*0401 and propose a model for bound CII peptide (261–273), the immunodominant peptide in human–transgenic mouse models of RA (Fugger et al., 1996; Rosloniec et al., 1997).

Results and Discussion

The Superantigen SEB Facilitated Growth of DR4/Peptide Crystals Ordered Well Enough for Structure Determination by X-Ray Crystallography

Although small crystals of HLA-DR4 (DRA*0101, DRB1*0401) complexed with the type II collagen peptide (1168–1180) could be grown from DR4 expressed in insect cells and although crystals of DR4 diffracting to 3.7Å resolution have been grown from DR4 purified from human lymphoblastoid cells (Gorga et al., 1991), it was necessary to add the superantigen SEB to DR4/CII (1168–1180) to grow crystals suitable for high-resolution (2.5Å) structural studies. The addition of SEB has also been used by Stallings and colleagues to grow DR4 crystals complexed with an influenza virus HA peptide (Stallings et al., 1996, Acta Crystallogr., abstract). The structure of DR4/CII(1168–1180)/SEB was determined by molecular replacement using coordinates from the DR1/SEB complex (Jardetzky et al., 1994) as a search model, from which the 16 residues differing between DR1 and DR4 were substituted by Ala (see Experimental Procedures). The final model includes all but 19 residues from two disordered loops of SEB (dotted lines in Figure 1) and the last residue of the peptide, for which the electron density is weak, suggesting that it is disordered (Figure 2).

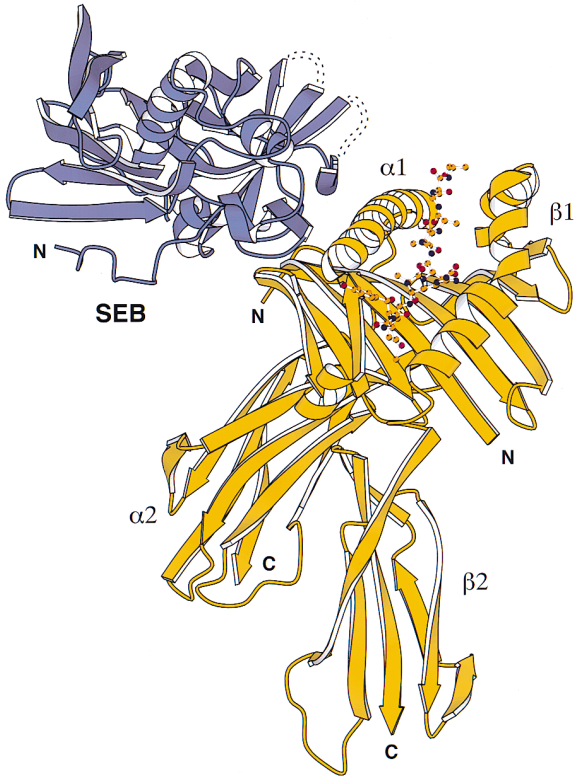


Figure 1. Ribbon Diagram of the HLA-DR4/CII(1168–1180)/SEB Complex

Peptide atoms are shown as balls and sticks. Two loops from SEB that are not visible in electron density maps are represented by dotted lines.

The Structure of the DR4/Collagen Peptide/SEB Complex Resembles Other MHC Class II/Peptide and MHC class II/Peptide/Superantigen Complexes

The overall structure of the SEB/DR4 complex is very similar to that of SEB/DR1, since DR1 and DR4 share the same α chain, to which SEB binds (Jardetzky et al., 1994). However, no dimer of αβ dimers is observed in the SEB/DR4 crystal (compare Figure 1 to Figure 1D in Jardetzky et al., 1994). The bound collagen peptide (QY**M**RADQAAGGLR; M = P1) is extended in a polyproline type II conformation (Stern et al., 1994a; Ghosh et al., 1995; Jardetzky et al., 1996) with the Met-1170, P1, occupying the major nonpolar pocket, first identified in DR1 (Brown et al., 1993; Stern et al., 1994a) (Figure 3A). As in other class II MHC/peptide structures, peptide side chains at P1, P4, P6, and P7 project into pockets in the peptide-binding groove (P9, which usually fits in a pocket is Gly, with no side chain in CII(1168–1180)) (Figure 3A). The arginine at P2 hydrogen-bonds to the carbonyl oxygen of β chain Thr-77.

There are 11 hydrogen bonds between main-chain peptide atoms and DR4, forming a network of hydrogen bonds to conserved class II residues that has been found in all class II MHC/peptide complexes to date

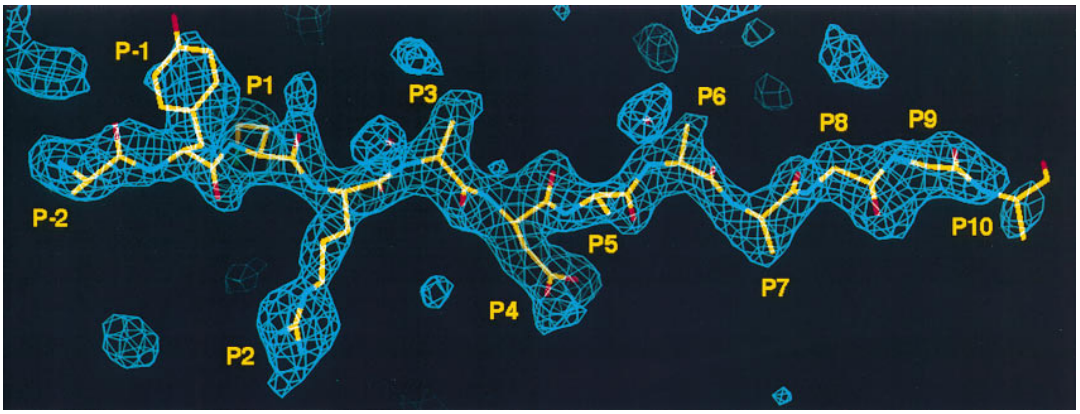
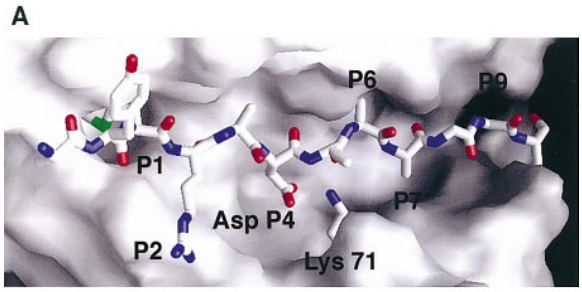


Figure 2. CII(1168–1180) Peptide in the HLA-DR4 Binding Site Overlaid with an Fo-Fc Omit Map Contoured at 2.2 σ

Collagen peptide atoms were omitted from electron density map calculation. Gln at P −2, Gln at P5, and Leu at P10 were modeled as Ala.

(Stern et al., 1994a; Stern and Wiley, 1994b; Ghosh et al., 1995; Fremont et al., 1996; Jardetzky et al., 1996) (Figure 3B). Two hydrogen bonds to the peptide main chain at P10, formed by Asp-β57 and Arg-α76 in class II MHC molecule binding clefts, are not observed in DR4/CII(1168–1180) because the peptide C-terminus is solvent exposed and appears to be partly disordered beyond the Gly at P9 in this crystal. Two peptide residues N-terminal to the P1 Met, the Tyr at P −1 and the backbone atoms for Gln at P −2, are well ordered (Figures 2 and 3A).

In the DR4/SEB crystal lattice, a second extensive interaction is observed between DR4 and SEB, in addition to that described in the DR1/SEB complex (Jardetzky et al., 1994). It is almost certainly not physiological, because the two class II molecules are oriented almost perpendicular to each other, an arrangement that would not be expected to occur on the cellular membrane. This interaction involves pairing of two β-strands in the C-terminal portion of SEB (residues 129–149) and a β-strand involving residues 179–189 of the HLA-DR4 β-domain, to form an intermolecular β sheet. These contacts appear to stabilize SEB in the DR4/SEB crystal lattice, allowing regions that appeared disordered in the DR1/SEB crystal to be defined here (see Experimental Procedures).



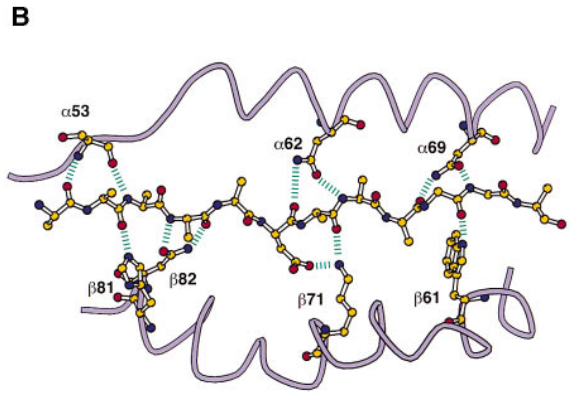


Figure 3. CII(1168–1180) Peptide in the HLA-DR4 Binding Site

(A) CII(1168–1180) peptide in the HLA-DR4 binding site. P1 Met (green) fits into a deep pocket. Asp P4 and Lys-β71 interact in pocket 4. P-2, P5, and P10 are displayed as Ala.

(B) Hydrogen bonds between CII(1168–1180) and HLA-DR4. Peptide side chains are shown as Ala, with the exception of Asp P4. Ten hydrogen bonds are made from the peptide main chain to conserved class II residues. Lys-β71 makes two other hydrogen bonds, one to the P4 Asp side chain and another to the carbonyl of P5.

The Interactions between the CII Peptide Side Chains and DR4 Match Motifs Derived from Phage-Display and Peptide Libraries

Peptides that bind to DRB1*0401 show amino acid preferences at six positions: P1, P2, P3, P4, P6, and P7. The sequence motif deduced from phage-display and synthetic peptide libraries for DRB1*0401 is: P1 large nonpolar, P2 Arg, P3 small nonpolar (Ala, Gly), P4 nonpolar, P6 Thr/Ser, and P7 aliphatic (Hammer et al., 1993, 1994, 1995). The CII(1169–1180)/DR4 complex shows the basis for these preferences.

The DR4 pocket that binds P1 Met (peptide residue 1170) is deep and nonpolar (Figure 3A). A Gly/Val dimorphism at DRβ86 in this pocket influences side chain specificity at P1 (Busch et al., 1991; Demotz et al., 1993; Newton-Nash and Eckels, 1993; Verreck et al., 1993). Glycine β86 as visualized here in DRB1*0401 and earlier in DR1 (Brown et al., 1993; Stern et al., 1994a; Jardetzky et al., 1996) allows the binding of large aromatic and nonpolar side chains, while Val as in DRB1*0402 and DRB1*0404 and visualized in DR3 (Ghosh et al., 1995) decreases the size of the pocket, excluding aromatic

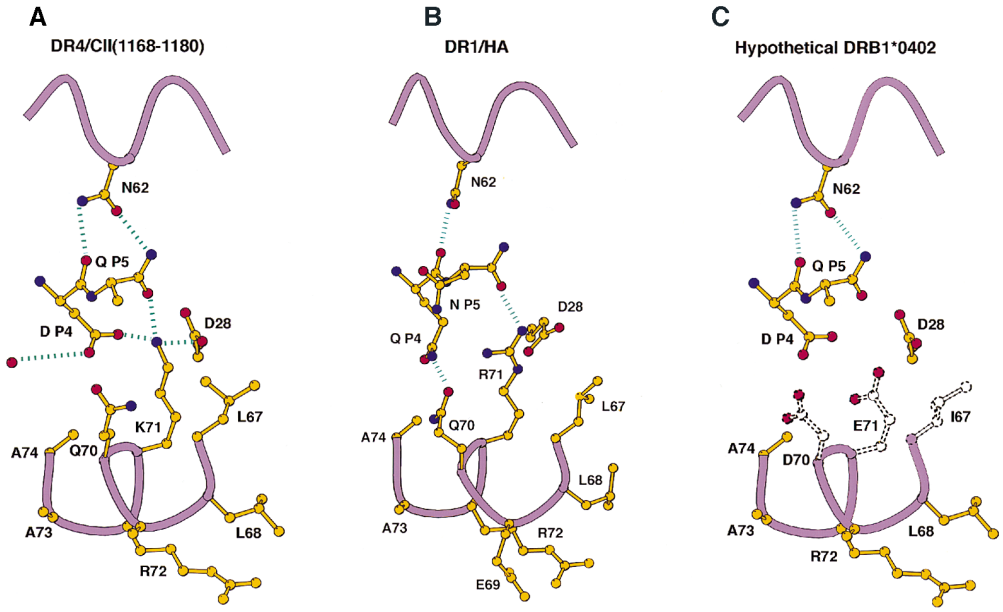


Figure 4. Interactions in the P4 Pockets of HLA-DR1, HLA-DR4 (DRB1*0401), and HLA-DR4 (DRB1*0402)

(A) Interatomic contacts in the P4 pocket of DR4/CII(1168–1180). Six hydrogen bonds are found in the pocket, including the one to a bound water molecule. Other residues forming the shared epitope are shown in yellow.

(B) Pocket 4 of the DR1/HA complex.

(C) Hypothetical model of pocket 4 of DRB1*0402, an HLA-DR4 allele that is not associated with the development of arthritis.

side chains. Substitutions at DRβ86 in DRB1*0401 markedly affect T cell recognition, probably by altering the peptide-binding characteristics of the molecule via its P1 pocket (Barber et al., 1991; Fu et al., 1995). Within DR4 subtypes, the DRβ86 dimorphism determines the P1 specificity but is not critical in the linkage to RA, since both residues occur in RA-associated and -nonassociated molecules (Hammer et al., 1995).

The side chain of P2 Arg extends across the binding cleft, so that it can be recognized by a T cell receptor (TCR), and its positively charged guanidinium group forms a hydrogen bond with the main-chain oxygen of Thr-β77 (Figure 3A). Hydrogen bonds to the same β77 carbonyl oxygen have been observed between P2 Thr and Asn residues in two mouse I-E^k/peptide complexes (Fremont et al., 1996). P2 Arg was also found frequently in peptides that bind to DRB1*0101 and DRB1*1101, other RA-associated alleles (Hammer et al., 1992, 1994).

The P3 Ala methyl group side chain is exposed to TCR contacts on a shelf-like pocket identical to those observed in DR1 and DR3, which accommodate the methylene groups of a solvent-exposed Lys in the DR1/HA complex (Stern et al., 1994a) and Met in the DR3/CLIP (Ghosh et al., 1995). In the DR4/CII complex, a water molecule occupies the approximate position where the ε-amino group of P3 Lys is bound in the DR1/HA complex (Stern et al., 1994a) and is hydrogen bonded to the main-chain carboxylate group of Gly-α58. Polymorphism in the DR pocket that binds P4 has been associated with both RA and pemphigus vulgaris (reviewed by Wucherpfennig and Strominger, 1995). In the DRB1*0401/CII(1168–1180) complex the Asp side chain of P4 is deep in the binding site (Figure 3A) and forms a salt-bridged hydrogen bond with the ε-amino group of Lys-β71 (Figure 4A). The latter residue, which is part of the α helix in the β domain, points directly into the peptide-binding site. Phage-display and peptide library studies revealed an increased frequency of nonpolar residues at P4 (Met, Ala, Val, and Leu) (Hammer et al., 1993), but charged residues binding in this pocket have been implicated in autoimmune diseases (Hammer et al., 1995; Wucherpfennig and Strominger, 1995) (see next section).

The glutamine side chain of P5 projects up and out of the peptide-binding site. Electron density for the side chain is broken (Figure 2), probably indicating disorder and complete solvent exposure. P5 may interact with Qβ70, which projects out of the binding site adjacent to P5. The P5 side chain is positioned projecting out of the peptide-binding site of MHC class II molecules at the same relative position as a P5 Tyr projects out of the MHC class I HLA-A2 binding site and into a pocket formed by hypervariable loops on a bound TCR, in the crystal structure of the ternary complex of A6-TCR/HTLV-1 Tax peptide/HLA-A2 (Garboczi et al., 1996a, 1996b).

The side chain of Ala at P6 is sequestered from solvent in a deep cleft (Figure 3A). Pocket 6 is deeper in DR4 than in DR1 because of the smaller side chain of β11, which is a valine in DR4 and a leucine in DR1. The side chain of Ala at P7, however, is only partly buried in a shallow pocket along the β chain helix (Figure 3A). A Tyr at β30 makes this pocket smaller in DR4 than DR1, which has a Ser as the β30 residue. The P8 Gly has only a hydrogen as a side chain in CII (1168–1180), but it is positioned so that other residues at P8 can project prominently into solvent.

Gly at P9 is positioned above a small, primarily nonpolar pocket that contains a bound water molecule. Leucine at P10 and arginine at P11 extend into solution

from the end of the peptide-binding site. The electron density in this region does not show the side chain positions, indicating that they are probably disordered (Figure 2).

The N-terminal peptide residues at positions P −2 (Gln) and P −1 (Tyr) form hydrogen bonds from the peptide main chain to conserved atoms on the DR4 (Figure 3B). The P −1 Tyr side chain projects upward out of the binding site (Figure 3A), but there is no electron density for the side chain of P −2, suggesting that it is disordered. Eleven water molecules that contact the bound peptide and adjacent residues in the binding site are included in the refined model. At the N-terminal end of the binding site, one water molecule is hydrogen bonded to the main-chain amide of the P −1 tyrosine, while another one forms a bridge between the Tyr hydroxyl group and OE1 of Glu-α55, at the top of the binding cleft. Another water molecule bridges DR4 residues Asn-62 (ND2) and Gly-58 (main-chain O), and the main-chain amide of P3 is bonded to a water molecule. OD1 of Asp at P4 also forms a contact with a tightly bound water (Figure 4A). The carbonyl oxygen of P6 Ala hydrogen-bonds with two closely placed water molecules, one of which is bonded to a third water that does not have direct contacts to protein or peptide atoms. The amide groups of Gly P7 and P8 hydrogen-bond with two water molecules, one of which bridges to OD1 of Asn-α69.

Peptide residues from P −2 to P8 are expected to be covered by the footprint of a TCR, based on the orientation and contact area seen in the class I/MHC/peptide interface (Garcia et al., 1996; Garboczi et al., 1996a). Peptide side chains at P −1, P2, and P5 project prominently toward the T cell, but P −2, P3, P4, P6, P7, P8, and P9 are also exposed and therefore may be recognized as well.

RA Susceptibility Correlates with Electrostatic Charge Differences in the P4 pockets of DRB1*0401 and DRB1*0402

A specific constellation of polymorphic residues on the α helix of the β chain adjacent to pocket 4, residues 67, 70, 71, and 74 (Figure 4), has been identified as creating an epitope shared by DR allelic products associated with RA (Silver and Goyert, 1985; Gregersen et al., 1987; Hiraiwa et al., 1990). In the DR4/CII(1168–1180) structure, Lys-β71 forms hydrogen bonds with both the main-chain carbonyl oxygen of P5 and the Asp side chain of P4 (Figure 4A), and Gln-β70 points out of the binding site toward the T cell. The main chain of the peptide is also held in place by hydrogen bonds from Asn-62, a conserved residue on the α helix across the binding groove (top in Figure 4A). In pocket 4 of the DR1/HA peptide complex, an Arg at β71 makes the same interaction as does the Lys in DR4 (Figure 4B), but Gln-β70 points down into the pocket to contact the P4 side chain.

The DRB1*0402 subtype, which is not associated with RA, has the acidic residues Asp and Glu at β70 and β71 (modeled in Figure 4C). The Glu at β71 in DRB1*0402 lacks a hydrogen bond donor and therefore would not be able to form a hydrogen bond to the main-chain carbonyl oxygen of P5 as did the Lys and Arg on DR4 and DR1 (compare Figure 4C to Figures 4A and 4B). Peptide-binding (Hammer et al., 1995; Rammensee et al., 1995) and site-directed mutation studies (Hammer et al., 1995) suggest that the electrostatic charge of β71 has a strong influence on peptide-binding specificity. Peptides in which P4 is positively charged bind to the DR4 subtypes that are not associated with RA (DRB1*0402) 100–1000 times better than to RA-associated subtypes (DRB1*0401, DRB1*0404), presumably because of the charge complementarity between a positive P4 and negative Glu at β71 in the former case and charge repulsion between a positive P4 and positive Lys at β71 in the latter case (Hammer et al., 1995). Peptides in which P4 is negatively charged, such as Asp P4 in DR4/CII (1168–1180), correspondingly bind better to RA-associated subtypes with a positively charged β71 (Figures 4A and 4B) than to the nonassociated subtypes with a negatively charged β71 (Figure 4C) (Hammer et al., 1995). Substitution of β71 to the oppositely charged residue in the DR4 subtypes reverses these preferences (Hammer et al., 1995). The observation in the DR4/CII(1168–1180) crystal structure of a salt-linked hydrogen bond forming from β71 to the negatively charged Asp at P4 is consistent with the earlier suggestion (and molecular model) that the selective binding of peptides by the P4 pocket may be a major part of the mechanism of MHC association to RA (Hammer et al., 1995).

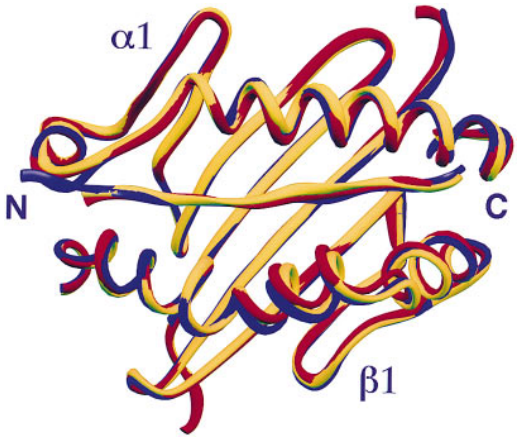


Figure 5. Comparison of DR1/HA, DR3/CLIP, and DR4/CII(1168–1180)

Superposition of DR1/HA (yellow), DR3/CLIP (blue), and DR4/CII (1168–1180) (red) shows small structural differences in the α helix of the β1 domains.

The location of the β chain α helix in DR4/CII(1168–1180) also differs from both DR1/HA and DR3/CLIP near residues β65 and β74 (Figure 5). In the DR3-CLIP structure, the peptide residue at position P4 is Ala, which has a small side chain, leading to a slightly "collapsed" pocket with the α-helical residues moved closer to the peptide (blue in Figure 5). An Asp side chain will not fit in the P4 pocket seen in DR3/CLIP structure, even though Asp is the preferred P4 residue for DR3 (Ghosh et al., 1995). In the DR4/CII (1168–1180) structure, pocket 4 accommodates the larger Asp side chain at P4 by a relative expansion of the pocket (red in Figure 5). Similarly, the P4 pocket in DR1/HA is larger (yellow in Figure

5), but the helical region between residues β65 and β74 is not exactly superimposable on the corresponding one in DR4/CII (1.0 Å root-mean-square deviation). These observations suggest that the exact positioning of this helix may vary according to the peptide side chain occupying the P4 pocket.

The Dominant Peptide Epitope in Collagen-Induced Arthritis of DR1 and DR4 Transgenic Mice (CII(261–273)) Can Be Modeled from the DR4/CII(1168–1180) Structure

Although the responsible antigen and the mechanism of DR linkage to increased susceptibility to RA is unknown in humans (e.g., Zanelli et al., 1995, 1996), there is evidence of both humoral and cellular responses to collagen II, which is a major component of joint cartilage (Stuart et al., 1983; Jasin, 1985; Watson et al., 1986; Londei et al., 1989; Terato et al., 1990). Transgenic mice expressing human DRB1*0401 and the human coreceptor CD4 (Fugger et al., 1994, 1996) as well as mice transgenic for chimeric DRB1*0101 (Rosloniec et al., 1997) harboring the DR1 peptide-binding site and mouse sequences in the class II MHC membrane-proximal CD4-interacting domains have been shown to generate a T cell response to immunization by CII. The same core CII peptide sequence (CII(263–270)) is immunodominant (Fugger et al., 1996; Rosloniec et al., 1997) in both transgenic mouse models.

We have modeled the CII(261–273) peptide (AG**F**K-G**E**QGPKGEP, P1 = **F**), which is recognized by the majority of CII reactive HLA-DRB1*0401–restricted T cells from the DR4 transgenic mouse model (Fugger et al., 1996) (Figure 6A) by analogy to the DR4/CII (1168–1180) structure (Figure 3A). The modeling is facilitated by the observation that the five peptides in the MHC class II structures determined to date, which include DR1, DR3, DR4, and I-E^k, have remarkably similar conformations (Figure 6B) (Stern et al., 1994a; Ghosh et al., 1995; Fremont et al., 1996; Jardetzky et al., 1996). In every case, P1, P4, P6, P7, and P9 are at least partially buried in pockets, while P −2, P −1, P2, P3, P5, P8, P10, and P11 are substantially exposed to solvent, with the side chains at P −1, P5, and P8 projecting directly up away from the MHC molecule surface (Figure 6B). As discussed earlier (Stern et al., 1994a; Jardetzky et al., 1996), the peptide main- chain conformations are regular and similar to those of polyproline type II. The main chains overlap almost perfectly from P −2 to P5 (Figure 6B), diverging after that as some of the peptides fit 1.0–2.0 Å deeper into the binding site from P6 to P10 (Ghosh et al., 1995).

The CII(261–273) peptide is expected to align with Phe-263 (P1) bound in the large nonpolar pocket 1, which was occupied by Met in the DR4/CII(1168–1180) complex, with Glu-266 (P4) positioned in pocket 4 to form a salt-bridged hydrogen bond to Lys-β71 as does P4 Asp in CII(1168–1170) (compare Figure 6A to Figure 3A) (Fugger et al., 1996). In addition, the Lys at position P2 could be hydrogen bonding to the main chain of Thr-β77, as is Arg at P2 in the CII(1168–1180) peptide (Figures 6A and 3A). All of the peptide main- chain hydrogen bonds to DR4 (Figure 3B) can be maintained in the model. P6 Gly replaces Ala of CII(1168–1180). P7 Pro fits in the shallow pocket 7, but Tyr-30 or P7 may move slightly in the real structure, as the fit is tight. P9 is Gly in both CII peptides. The side chains at P −1 and P3, expected to extend into solution, are only the hydrogens of Gly residues, so that TCRs could fit very tightly to the MHC molecule at these positions. In contrast, P5 Gln and P8 Lys are expected, by analogy to previous peptides (Figure 6B), to extend prominently into solution and be contacted by TCRs. Although the DR4/CII(261–273) complex is only a hypothetical model, the analogy to CII(1168–1180) is so close and the mode of peptide binding to MHC class II molecules so conserved, that the model is probably accurate enough for the planning of experiments. The differences between the model and the actual structure are probably less than the differences in peptide structure between the Tax 9-mer peptide bound to HLA-A2 and the same peptide in the TCR/Tax/HLA-A2 crystal structure, where rearrangements of two peptide residues of up to 4Å were observed following TCR binding (Garboczi et al., 1996a).

Experimental Procedures

Expression

The expression vectors for soluble DR4 were made by reverse transcription–polymerase chain reaction (PCR) of mRNA isolated from the DRB1*0401-expressing EB virus-transformed B cell line, Priess. The PCR primers amplified sequences encoding amino acids 1–217 of the α chain and 1–227 of the β chain. The PCR products were subcloned into the EcoRI/BamHI sites of the pRmHA-3 expression vector (Bunch et al., 1988). Stable transformants of Drosophila S2 cells were generated by cotransformation with pUChsneo (Stellar and Pirrotta, 1985) and selection with G418. High-secreting single cell–derived clones were isolated by limiting dilution culture. Cells were grown in spinner flasks in ExCell 401 medium (Sigma) supplemented with 1% fetal bovine serum (Sigma) to a density of 5×10^6 cells/ml, after which expression was induced by addition of 1 mM $CuSO_4$. Cells were allowed to incubate for another 6 days before harvesting, after which they were centrifuged at 4°C for cellular debris removal.

Purification and Crystallization

Recombinant DR4 was purified by immunoaffinity chromatography; specifically, the supernatant from the harvested cells was loaded onto a Protein A column and subsequently onto a monoclonal antibody LB3.1 column. Bound protein was eluted with 50 mM CAPS (pH 11.5) and neutralized with 200 mM phosphate (pH 6.0). Binding of peptide QYMRADQAAGGLR (CII 1168–1180) to the eluted HLA-DR4 was performed at 37°C for 48 hr in the presence of 5% glycerol. After incubation, the DR4-peptide complex was purified by gel filtration chromatography using a Superdex 200 column in 50 mM Tris-Cl (pH 8.0) + 0.1 M NaCl; the eluted peak was exchanged into 25 mM Tris-Cl (pH 8.0) buffer and concentrated to 8 mg/ml in Centricon 10 concentrators (Amicon). SEB (Sigma; suspended in the latter buffer) was added in a 1.5:1 molar ratio prior to crystallization. Crystals of the DR4-CII(1168–1180)/SEB complex were obtained at 25°C in 0.1 M sodium acetate (pH 4.6), 9%–14% polyethylene glycol (PEG) 4K, 0.2 M ammonium acetate, as well as in 0.1 M glycine (pH 3.85), 9%–14% PEG 4K, 0.1 M ammonium sulfate. Crystals appeared in 2 days and grew to dimensions of $0.5 \times 0.5 \times 0.2$ mm.

Data Collection and Processing

Data were collected at the Cornell High Energy Synchrotron Source (CHESS) using a CCD detection system (λ = 0.91 Å). A crystal grown in the sodium acetate condition was introduced into a solution of 0.1 M ammonium sulfate, 15% PEG 4K, and 25% glycerol prior to mounting and flash-cooling; 1.0° oscillation frames were collected, and all data were indexed and integrated using DENZO (Otwinowski,

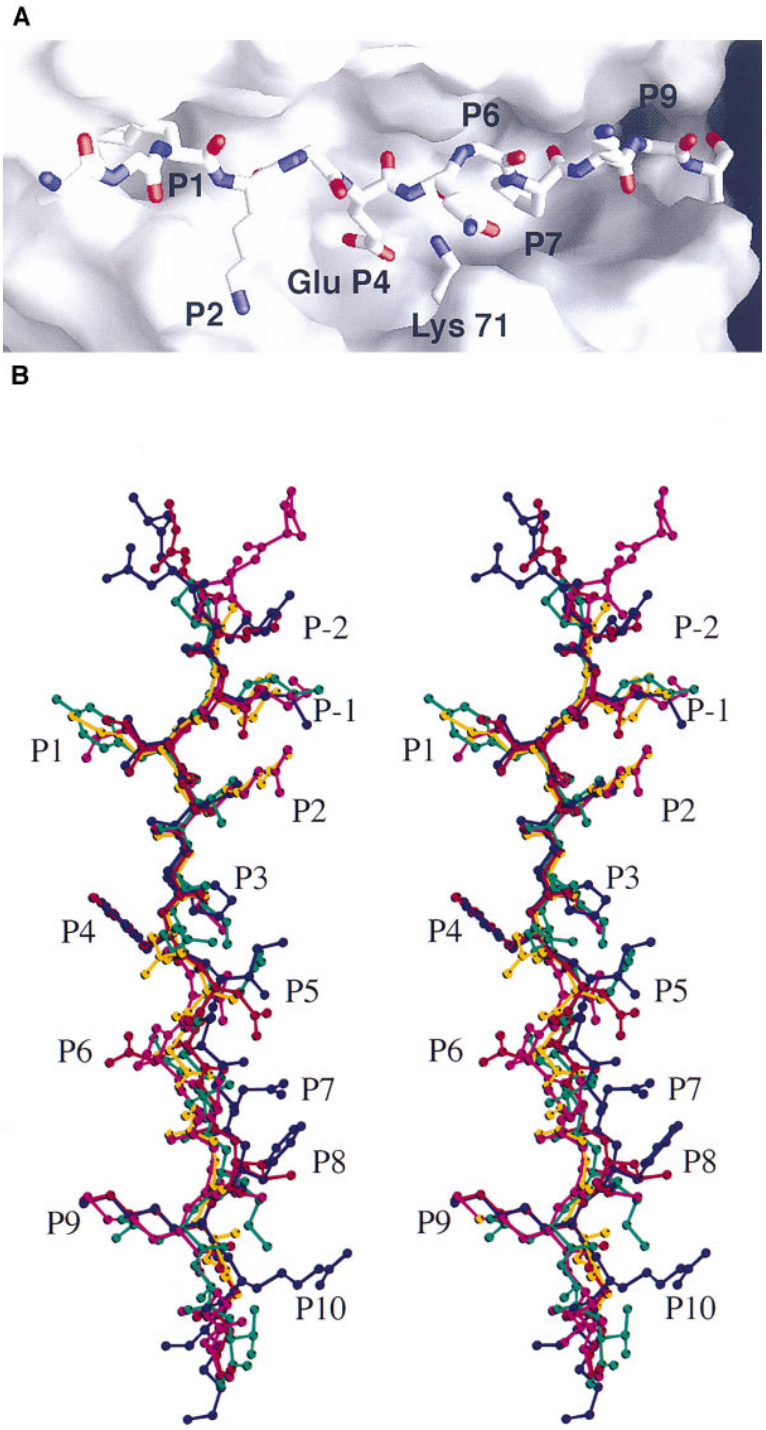


Figure 6. Conformations of Peptides in the Binding Sites of MHC Class II Molecules

(A) Hypothetical model of CII(261–273) peptide in the HLA-DR4 binding cleft. Phe-263 fits into the P1 pocket, and Glu (P4) hydrogen bonds to β71.

(B) The conformations of five peptides observed in complexes with human and murine class II MHC molecules: DR1/HA (green), DR3/CLIP (purple), DR4/CII(1168–1180) (yellow), Ie^k/HB (red), and Ie^k/HSP70 (blue). The MHC class II molecules were superimposed to reveal the superposition of the peptides. View is approximately with MHC to the left and TCR to the right.

1993) and scaled using SCALEPACK (CCP4). Data collected at CHESS are 87.7% complete to a resolution of 2.5 Å. The DR4-CII(1168–1180)-SEB complex crystallizes in orthorhombic space group $P2_12_12_1$, with unit cell parameters a = 78.52 Å, b = 100.28 Å, c = 100.86 Å, and $\alpha = \beta = \gamma = 90°$.

Molecular Replacement

A molecular replacement solution was identified with XPLOR (Brünger, 1992) using the DR1/SEB complex (Jardetzky et al., 1994) as a search model, for which 16 residues differing between DR1 and DR4 were substituted by Ala. After cross-rotation (12–3.0 Å) and translation (12–3.0 Å) procedures, the top solution (R_{cryst} 42%) was 9 σ higher than the next solution. The asymmetric unit of the crystal contains a single αβ heterodimer complexed to a single SEB molecule; neither of the HLA class II heterodimers observed previously (Brown et al., 1993; Jardetzky et al., 1994; Stern et al., 1994a; Ghosh et al., 1995; Fremont et al., 1996) was identified in this lattice.

Model Building and Refinement

A random 10% of reflections were omitted from refinement for a calculation of R_{free} (Brünger, 1992) prior to refinement. To improve

Table 2. Data Collection and Refinement Statistics

Resolution limit (Å)	2.5	
Mosaicity (°)	0.70	
R_{merge} (%)	8.1 (21.0)	
Unique reflections	24221	
Total observations	178959	
Completeness	87.7 (83.7)	
Refinement		
R_{free} (%)	28.0	
R_{cryst} (%)	22.9	
	Residue	Average B
Protein	598	20.4
Water	58	24.8
Root-mean-square deviations		
Bonds (Å)	Angles (°)	B factors (Å^2 bonded)
0.007	1.5	2.1

$R_{merge} = \Sigma \mid I - <I> \mid / \Sigma I$, where I is the observed intensity and $<I>$ is the average intensity of multiple observations of symmetry-related reflections.

$R = \Sigma \mid\mid Fo \mid - \mid Fc \mid\mid / \Sigma \mid Fo \mid$, where R_{free} is calculated for a randomly chosen 10% of reflections, R_{cryst} is calculated for the remaining 90% of reflections ($F > 2.0$) used for structure refinement.

Numbers in parentheses indicate the specific value in the resolution shell from 2.5 to 2.6 Å.

the geometry of the SEB model, the coordinates of SEC (1SE2; (Swaminathan et al., 1992), with residues that differed from SEB mutated to Ala, was rigid body–refined onto DR4-SEB; the newly generated DR4-superantigen complex structure had an initial R_{free} of 36.5% and R_{cryst} of 35.6%. Residues specific to DR4 and SEB were built into 2Fo-Fc and Fo-Fc electron density maps (12–3.0 Å) using O (Jones et al., 1991). The resolution limit was increased gradually to 2.5 Å through subsequent manual model building and refinement cycles, which included positional and restrained atomic B factor refinement, geometric regularization, and group B factor refinement. The peptide was modeled into clear, uninterrupted Fo-Fc density after R_{free} had dropped below 32% (12–2.7 Å). After an additional drop of 1% in R_{free}, water molecules within 2.6–3.4 Å of a hydrogen bonding donor or acceptor were built into Fo-Fc electron density $>3.5\ \sigma$. The final model contains 58 waters. 2Fo-Fc electron density maps show clear electron density for all residues of the α subunit; two breaks in the main-chain density of the β subunit occur in loop region of the $\beta 2$ domain (loops between residues 105 and 110, as well as 163 and 168). In the SEB molecule, two breaks in the main-chain density occur in surface loop regions, between residue 54 and 61 and 97 and 110. Comparisons of the DR4/SEB structure with the DR1/SEB complex (Jardetzky et al., 1994) reveal that the loops encompassing SEB residues 122–126 and 176–182, which were not observable in the latter structure, are clearly defined by electron density in the DR4/SEB structure. Also, the loop formed by residues 150–155 is slightly different. One N-acetyl glucosamine monosaccharide residue was modeled at each DR4 N-linked glycosylation site, $\alpha 78$N and $\alpha 118$N. Data collection and refinement statistics are shown in Table 2. Coordinates will be deposited in the Brookhaven Protein Databank and are available prerelease from A. D. (e-mail: dessen@xtal22.tch.harvard.edu).

Acknowledgments

We thank Peg Pietras for excellent technical assistance with tissue culture; N. Sinitskaya for peptide synthesis; A. Planchart, P. Rosenthal, and the staff at CHESS for assistance with data collection, and D. Fremont for sharing coordinates of the MHC class II I^k complexes. This work was supported by the National Institutes of Health. D. C. W. is an investigator of the Howard Hughes Medical Institute.

Received June 30, 1997; revised September 4, 1997.

References

Banerjee, S., Luthra, H.S., Moore, S.B., and O'Fallon, W.M. (1988). Serum IgG anti-native type II collagen antibodies in rheumatoid arthritis: association with HLA DR4 and lack of clinical correlation. Clin. Exp. Rheumatol. *6*, 373–380.

Barber, L.D., Bal, V., Lamb, J.R., O'Hehir, R.E., Yendle, J., Hancock, R.J.T., and Lechler, R.I. (1991). Contribution of T-cell receptor-contacting and peptide-binding residues of the class II molecule HLA-DR4 to serologic and antigen-specific T-cell recognition. Human Immunol. *32*, 110–118.

Brown, J.H., Jardetzky, T.S., Gorga, J.C., Stern, L.J., Urban, R.G., Strominger, J.L., and Wiley, D.C. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature *364*, 33–39.

Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472–475.

Bunch, T.A., Grunblat, Y., and Goldstein, L.S.B. (1988). Characterization and use of the *Drosophila* metallothionein promoter in cultured *Drosophila melanogaster* cells. Nucleic Acids Res. *16*, 1043–1061.

Busch, R., Hill, C.M., Hayball, J.D., Lamb, J.R., and Rothbard, J.B. (1991). Effect of natural polymorphism at residue 86 of the HLA-DR beta chain on peptide binding. J. Immunol. *147*, 1292–1298.

Demotz, S., Barbey, C., Corradin, G., Amoroso, A., and Lanzavecchia, A. (1993). The set of naturally processed peptides displayed by DR molecules is tuned by polymorphism of residue 86. Eur. J. Immunol. *23*, 425–432.

Feldman, M., Brennan, F.M., and Maini, R.V. (1996). Rheumatoid arthritis. Cell *85*, 307–310.

Fremont, D.H., Hendrickson, W.A., Marrack, P., and Kappler, J. (1996). Structures of an MHC class II molecule with covalently bound single peptides. Science *272*, 1001–1004.

Fu, X.T., Bono, C.P., Woulfe, S.L., Swearingen, C., Summers, N.L., Sinigaglia, F., Sette, A., Schwartz, B.D., and Karr, R.W. (1995). Pocket 4 of the HLA-DR(alpha,beta 1*0401) molecule is a major determinant of T cells recognition of peptide. J. Exp. Med. *181*, 915–926.

Fugger, L., Michie, S.A., Rulifson, I., Lock, C.B., and McDevitt, G.S. (1994). Expression of HLA-DR4 and human CD4 transgenes in mice determines the variable region beta-chain T-cell repertoire and mediates an HLA-DR-restricted immune response. Proc. Natl. Acad. Sci. USA *91*, 6151–6155.

Fugger, L., Rothbard, J.B., and Sonderstrup-McDevitt, G. (1996). Specificity of an HLA-DRB1*0401-restricted T cell response to type II collagen. Eur. J. Immunol. *26*, 928–933.

Garboczi, D.N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W.E., and Wiley, D.C. (1996a). Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature *384*, 134–141.

Garboczi, D.N., Utz, U., Ghosh, P., Seth, A., Kim, J., van Tienhoven, A.E., Biddison, W.E., and Wiley, D.C. (1996b). Assembly, specific binding, and crystallization of a human TCR-$\alpha\beta$ with an antigenic Tax peptide from human T lymphotropic virus type 1 and the class I MHC molecule HLA-A2. J. Immunol. *157*, 5403–5410.

Garcia, K.C., Degano, M., Stanfield, R.L., Brunmark, A., Jackson, M.R., Peterson, P.A., Teyton, L., and Wilson, I.A. (1996). An $\alpha\beta$ T cell receptor structure at 2.5Å and its orientation in the TCR-MHC complex. Science *274*, 209–219.

Ghosh, P., Amaya, M., Mellins, E., and Wiley, D.C. (1995). The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. Nature *378*, 457–462.

Gorga, J.C., Brown, J.H., Jardetzky, T., Wiley, D.C., and Strominger, J.L. (1991). Crystallization of HLA-DR antigens. Res. Immunol. *142*, 401–407.

Gregersen, P.K., Silver, J., and Winchester, R.J. (1987). The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum. *30*, 1205–1213.

Hammer, J., Takacs, B., and Sinigaglia, F. (1992). Identification of a motif for HLA-Dr1 binding peptides using M13 display libraries. J. Exp. Med. *176*, 1007–1013.

Hammer, J., Valsasnini, P., Tolba, K., Bolin, D., Higelin, J., Takacs,

B. and Sinigaglia, F. (1993). Promiscuous and allele-specific anchors in HLA-DR-binding peptides. Cell *74*, 197–203.

Hammer, J., Belunis, C., Bolin, D., Papadopoulos, J., Walsky, R., Higelin, J., Danho, W., Sinigaglia, F., and Nagy, Z.A. (1994). High-affinity binding of short peptides to major histocompatibility complex class II molecules by anchor combinations. Proc. Natl. Acad. Sci. USA *91*, 4456–4460.

Hammer, J., Gallazzi, F., Bono, E., Karr, R.W., Guenot, J., Valsasnini, P., Nagy, Z.A., and Sinigaglia, F. (1995). Peptide binding specificity of HLA-DR4 molecules: correlation with rheumatoid arthritis association. J. Exp. Med. *181*, 1847–1855.

Hiraiwa, A., Yamanaka, K., Kwok, W.W., Mickelson, E.M., Masewicz, S., Hansen, J.A., Radka, S.F., and Nepom, G.T. (1990). Structural requirements for recognition of the HLA-Dw14 class II epitope: a key HLA determinant associated with rheumatoid arthritis. Proc. Natl. Acad. Sci. USA *87*, 8051–8055.

Jardetzky, T.S., Brown, J.H., Gorga, J.C., Stern, L.J., Urban, R.G., Chi, Y.I., Stauffacher, C., Strominger, J.L., and Wiley, D.C. (1994). Three-dimensional structure of a human class II histocompatibility molecule complexed with superantigen. Nature *368*, 711–718.

Jardetzky, T.S., Brown, J.H., Gorga, J.C., Stern, L.J., Urban, R.G., Strominger, J.L., and Wiley, D.C. (1996). Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. Proc. Natl. Acad. Sci. USA *93*, 734–738.

Jasin, H.E. (1985). Autoantibody specificities of immune complexes sequestered in articular cartilage of patients with rheumatoid arthritis and osteoarthritis. Arthritis Rheum. *28*, 241–248.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Khare, S.D., Krco, C.J., Griffiths, M.M., Luthra, H.S., and David, C.S. (1995). Oral administration of an immunodominant human collagen peptide modulates collagen-induced arthritis. J. Immunol. *155*, 3653–3659.

Lanchbury, J.S., Jaeger, E.E., Sansom, D.M., Hall, M.A., Wordsworth, P., Stedeford, J., Bell, J.I., and Panayi, G.S. (1991). Strong primary selection for the Dw4 subtype of DR4 accounts for the HLA-DQw7 association with Felty's syndrome. Hum. Immunol. *32*, 56–64.

Londei, M., Savill, C.M., Verhoef, A., Brennan, F., Leech, Z.A., Duance, V., Maini, R.N., and Feldman, M. (1989). Persistence of collagen type II-specific T-cell clones in the synovial membrane of a patient with rheumatoid arthritis. Proc. Natl. Acad. Sci. USA *86*, 636–640.

Mengle-Gaw, L., Conner, S., McDevitt, H.O., and Fathman, C.G. (1984). Gene conversion between murine class II major histocompatibility complex loci: functional and molecular evidence from the bm 12 mutant. J. Exp. Med. *160*, 1184–1194.

Nepom, G.T., and Erlich, H. (1991). MHC class-II molecules and autoimmunity. Annu. Rev. Immunol. *9*, 493–525.

Nepom, G.T., Byers, P., Seyfried, C., Healey, L.A., Wilske, K.R., Stage, D., and Nepom, B.S. (1989). HLA genes associated with rheumatoid arthritis: identification of susceptibility alleles using specific oligonucleotide probes. Arthritis Rheum. *32*, 15–21.

Newton-Nash, D.K., and Eckels, D.D. (1993). Differential effect of polymorphism at HLA-DR1 beta-chain positions 85 and 86 on binding and recognition of DR1-restricted antigenic peptides. J. Immunol. *150*, 1813–1821.

Otwinowski, Z. (1993). In Proceedings of CCP4 Study Weekend: Data Collection and Processing, L. Sawyer, N. Isaacs, S. Bailey, eds. (Warrington, UK: SERC Daresbury Laboratory), pp. 56–62.

Rammensee, H.G., Friede, T., and Stevanoviic, S. (1995). MHC ligands and peptide motifs: first listing. Immunogenetics *41*, 178–228.

Rosloniec, E.F., Brand, D.D., Myers, L.K., Whittington, K.B., Gumanovskaya, M., Zaller, D.M., Woods, A., Altmann, D.M., Stuart, J.M., and Kang, A.H. (1997). An HLA-DR1 transgene confers susceptibility to collagen-induced arthritis elicited with human type II collagen. J. Exp. Med. *185*, 1113–1122.

Signorelli, K.L., Watts, L.M., and Lambert, L.E. (1995). The importance of DR4Dw4 beta chain residues 70, 71, and 86 in peptide binding and T cell recognition. Cell. Immunol. *162*, 217–224.

Silver, J., and Goyert, S.M. (1985). Epitopes Are the Functional Units of Ia Molecules and Form the Molecular Basis for Disease Susceptibility (Berlin: Springer-Verlag).

Steere, A.C., Dwyer, E., and Winchester, R. (1990). Association of chronic Lyme arthritis with HLA-DR4 and HLA-DR2 alleles. N. Engl. J. Med. *323*, 219–223.

Stellar, H., and Pirrotta, V. (1985). A transposable P vector that confers selectable G418 resistance to Drosophila larvae. EMBO J. *4*, 167–171.

Stern, L.J., and Wiley, D.C. (1994b). Antigenic peptide binding by class I and class II histocompatibility proteins. Structure *2*, 245–251.

Stern, L.J., Brown, J.H., Jardetzky, T.S., Gorga, J.C., Urban, R.G., Strominger, J.L., and Wiley, D.C. (1994a). Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. Nature *368*, 215–221.

Stuart, J.M., Huffstutter, E.H., Townes, A.S., and Kang, A.H. (1983). Incidence and specificity of antibodies to type I, II, III, IV, and V collagen in rheumatoid arthritis and other rheumatic diseases as measured by ^{125}L-radioimmunoassay. Arthritis Rheum. *26*, 832–840.

Swaminathan, S., Furey, W., Pletcher, J., and Sax, M. (1992). Crystal structure of staphylococcal enterotoxin B, a superantigen. Nature *359*, 801–806.

Terato, K., Shimozuru, Y., Katayama, K., Takemitsu, Y., Yamashita, I., Miyatsu, M., Fujii, K., Sagara, M., Kobayashi, S., and Goto, M., et al. (1990). Specificity of antibodies to type II collagen in rheumatoid arthritis. Arthritis Rheum. *33*, 1493–1500.

Tiwari, J., and Terasaki, P. (1985). HLA and Disease Association. (New York: Springer-Verlag).

Todd, J.A., Acha-Orbea, H., Bell, J.I., Chao, N., Fronek, Z., Jacob, C.O., McDermott, M., Sinha, A.A., Timmerman, L., Steinman, L., et al. (1988). A molecular basis for MHC class II–associated autoimmunity [published erratum appears in Science, *241*, 888, 1988]. Science *240*, 1003–1009.

Verreck, F.A., Termijtelen, A., and Koning, F. (1993). HLA-DR beta chain residue 86 controls DR alpha beta dimer stability. Eur. J. Immunol. *23*, 1346–1350.

Watson, W., Cremer, M., Wooley, P., and Townes, A. (1986). Assessment of the potential pathogenicity of type II collagen autoantibodies in patients with rheumatoid arthritis. Arthritis Rheum. *29*, 1316–1321.

Woulfe, S.L., Bono, C.P., Zacheis, M.L., Kirschmann, D.A., Baudino, T.A., Swearingen, C., Karr, R.W., and Schwartz, B.D. (1995). Negatively charged residues interacting with the p4 pocket confer binding specificity to DRB1*0401. Arthritis Rheum. *38*, 1744–1753.

Wucherpfennig, K.W., and Strominger, J.L. (1995). Selective binding of self peptides to disease-associated major histocompatibility complex (MHC) molecules: a mechanism for MHC-linked susceptibility to human autoimmune diseases. J. Exp. Med. *181*, 1597–1601.

Zanelli, E., Gonzalez-Gay, M.A., and David, C.S. (1995). Could HLA-DRB1 be the protective locus in rheumatoid arthritis? Immunol. Today *16*, 274–278.

Zanelli, E., Krco, C.J., Baisch, J.M., Cheng, S., and David, C.S. (1996). Immune response of HLA-DQ8 transgenic mice to peptides from the third hypervariable region of HLA-DRB1 correlates with predisposition to rheumatoid arthritis. Proc. Natl. Acad. Sci. USA *93*, 1814–1819.